

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 20, 1999, 15:16:47 : Search time 505.23 Seconds

(Without alignments)  
11922.305 Million cell updates/sec

Title: US-09-126-945-1

Perfect score: 1894  
Sequence: 1 gtccgactctctcccccagcac.....ataaagatctactagagactg 1894

Scoring table: OLIGODX\_NUC

Searched: 679419 seqs, 1590154680 residues

Database:

GenEmbl:\*  
1: gb\_bal:\*  
2: gb\_bal:\*  
3: gb\_bal:\*  
4: gb\_bal:\*  
5: gb\_bal:\*  
6: gb\_bal:\*  
7: gb\_bal:\*  
8: gb\_bal:\*  
9: gb\_bal:\*  
10: gb\_bal:\*  
11: gb\_bal:\*  
12: gb\_bal:\*  
13: gb\_bal:\*  
14: gb\_bal:\*  
15: gb\_bal:\*  
16: gb\_bal:\*  
17: gb\_bal:\*  
18: gb\_bal:\*  
19: gb\_bal:\*  
20: gb\_bal:\*  
21: gb\_bal:\*  
22: gb\_bal:\*  
23: gb\_bal:\*  
24: gb\_bal:\*  
25: gb\_bal:\*  
26: gb\_bal:\*  
27: gb\_bal:\*  
28: gb\_bal:\*  
29: gb\_bal:\*  
30: gb\_bal:\*  
31: gb\_bal:\*  
32: gb\_bal:\*  
33: gb\_bal:\*  
34: gb\_bal:\*  
35: gb\_bal:\*  
36: gb\_bal:\*  
37: gb\_bal:\*  
38: gb\_bal:\*  
39: gb\_bal:\*  
40: gb\_bal:\*  
41: gb\_bal:\*  
42: gb\_bal:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1894	100.0	1894 11 AF071538	AF071538 Homo sapi

2	38	2.0	1704	12	AB019436	AB019436 Mus muscu
3	30	1.6	848	4	AF075706	AF075706 Gallus ga
4	29	1.5	2508	4	AF075708	AF075708 Gallus ga
5	28	1.5	60056	35	AC007818	AC007818 Drosophila
6	28	1.5	362	36	DR02784A	M88474 Drosophila
7	25	1.3	1403	10	HS0194E3	Y08979 H.sapiens F
8	25	1.3	1901	10	HS0194E3	Y08979 H.sapiens m
9	24	1.3	480	36	DR02784A	M88474 Drosophila
10	24	1.3	85095	37	AC004573	AC004573 Drosophila
11	22	1.2	115835	9	HS28C20	AL031119 Human DNA
12	22	1.2	2658	12	MMNETRN	Z32815 M.musculus
13	22	1.2	2211	12	MUSTRANONC	L119953 Mus musculus
14	22	1.2	1526	12	SR2864	S82864 Elk-3-Ets t
15	22	1.2	113899	36	PMAL3P4	AL0088970 Plasmodiu
16	21	1.1	1131453	34	HS0898H1	283851 Homo sapien
17	20	1.1	128598	1	D90912	D90912 Synchocyst
18	20	1.1	473	5	E06024	E06024 cDNA encodl
19	20	1.1	2073	5	E06025	E06025 cDNA encodl
20	20	1.1	36742	9	D87002	D87002 Human (lamb
21	20	1.1	2064	9	HUMELAF	D12765 Human mRNA
22	20	1.1	152141	10	HS007000	U07000 Human break
23	20	1.1	2657	10	HS015635	U15635 Human ets d
24	20	1.1	2333	10	HS018018	U18018 Human E1A e
25	20	1.1	35687	11	AC000051	AC000051 Homo sapi
26	20	1.1	34539	11	AC002051	AC002051 Homo sapi
27	20	1.1	151111	11	AC002308	AC002308 Homo sapi
28	20	1.1	36921	11	AC004185	AC004185 Homo sapi
29	20	1.1	112220	11	AC005373	AC005373 Homo sapi
30	20	1.1	37139	11	AC006047	AC006047 Homo sapi
31	20	1.1	111084	11	AC006486	AC006486 Homo sapi
32	20	1.1	2642	11	HSETSF5	AF095890 Homo sapi
33	20	1.1	1046	12	MUSETSRPA	L10427 Mus musculu
34	20	1.1	126983	34	AC005283	AC005283 Homo sapi
35	20	1.1	236822	34	D84394	D84394 Homo sapien
36	20	1.1	111930	34	HSB7LIC4	AL078476 Homo sapi
37	20	1.1	269150	34	HS0194E24	AL050317 Homo sapi
38	20	1.1	160979	35	AC007664	AC007664 Homo sapi
39	20	1.1	102353	35	AC007708	AC007708 Homo sapi
40	20	1.1	58839	35	AC007821	AC007821 Drosophila
41	20	1.1	1832	37	AF106964	AF106964 Strongylo
42	20	1.1	43881	42	AC002054	AC002054 Homo sapi
43	20	1.1	155313	42	AC006453	AC006453 Homo sapi
44	19	1.0	35516	1	MTCC373	273419 Mycobacteri
45	19	1.0	40790	1	MTCC493	295844 Mycobacteri

## ALIGNMENTS

RESULT	1	LOCUS	AF071538	DEFINITION	Homo sapiens Ets transcription factor PDEF (PDEF) mRNA, complete cds.	ACCESSION	AF071538	NID	94007417	VERSION	AF071538.1	GI:4007417	ORGANISM	human.	REFERENCE	1 (bases 1 to 1894)	AUTHORS	Oettingen,P., Finger,E., Akbarali,Y., Thamrongsak,U., Dube,A., Boltao,J., Brown,L., Kunsch,C. and Libermann,T.A.	TITLE	Isolation and characterization of a novel prostate epithelium-specific Ets transcription factor, PDEF	JOURNAL	Unpublished	REFERENCE	2 (bases 1 to 1894)	AUTHORS	Libermann,T.A., Oettingen,P., Kunsch,C. and Finger,E.	TITLE	Direct Submission	Submitted (10-JUN-1998)	Medicine, Beth Israel Deaconess Medical Center, 330 Brookline Ave., Boston, MA 02215, USA
RESULT	1	LOCUS	AF071538	DEFINITION	Homo sapiens Ets transcription factor PDEF (PDEF) mRNA, complete cds.	ACCESSION	AF071538	NID	94007417	VERSION	AF071538.1	GI:4007417	ORGANISM	human.	REFERENCE	1 (bases 1 to 1894)	AUTHORS	Oettingen,P., Finger,E., Akbarali,Y., Thamrongsak,U., Dube,A., Boltao,J., Brown,L., Kunsch,C. and Libermann,T.A.	TITLE	Isolation and characterization of a novel prostate epithelium-specific Ets transcription factor, PDEF	JOURNAL	Unpublished	REFERENCE	2 (bases 1 to 1894)	AUTHORS	Libermann,T.A., Oettingen,P., Kunsch,C. and Finger,E.	TITLE	Direct Submission	Submitted (10-JUN-1998)	Medicine, Beth Israel Deaconess Medical Center, 330 Brookline Ave., Boston, MA 02215, USA

FEATURES  
source 1.1894 Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/map="6p21.3"  
/tissue\_type="prostate"  
1.1894  
/gene="PDEF"  
416..1423  
/gene="PDEF"  
/note="prostate epithelium-specific"  
/codon\_start=1  
/product="Ets transcription factor PDEF"  
/protein\_id="AAC95296.1"  
/db\_xref="pid:94007418"  
/db\_xref="gi:4007418"  
/translation="MGSASPGLSVSPSHLLPPDVSRTGLEKRAAGAVGLERRDMS  
PSPATPBOGLSAFLSYFDMLYPEDSSMAKAGASREPPPEPOCPYIDSOAFA  
GSLDVPGLTLEHSLDVOVMVGEVLKDIETACKLNTADPMSPSNOKXLT  
WTEHQYRLPMPGKAFOELAGKELCAMSEORROSPGLGDVLAHLIDINKSAMNKR  
TSPGAIHYCASTSESWTDSVDSGSPRIHLMQPLKELLKPHSTGRIRIWLNKK  
GIFKEDSAQVARKLMTGRNPRMAYIDLSRSINQYTKGILIRKPDISOHLVYQFVAP  
1"

BASE COUNT 368 a 653 c 571 g 302 t  
ORIGIN

Query Match 100.0%; Score 1894; DB 11; Length 1894;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1894; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gctcgaactctcccaagcaatctctgaactcgcgtgtccacactgccccacagacc 60  
Db 1 gctcgaactctcccaagcaatctctgaactcgcgtgtccacactgccccacagacc 60  
QY 61 agtctccaagcctgtctcgaactcctcctcgaagccccctcaggttggccttgccagcgtg 120  
Db 61 agtctccaagcctgtctcgaactcctcctcgaagccccctcaggttggccttgccagcgtg 120  
QY 61 agtctccaagcctgtctcgaactcctcctcgaagccccctcaggttggccttgccagcgtg 120  
Db 61 agtctccaagcctgtctcgaactcctcctcgaagccccctcaggttggccttgccagcgtg 120  
QY 121 ccagcagcagcagccttggtcgtggttaaggagactccctacagcagcagccttgagacc 180  
Db 121 ccagcagcagcagccttggtcgtggttaaggagactccctacagcagcagccttgagacc 180  
QY 121 ccagcagcagcagccttggtcgtggttaaggagactccctacagcagcagccttgagacc 180  
Db 121 ccagcagcagcagccttggtcgtggttaaggagactccctacagcagcagccttgagacc 180  
QY 181 tcaagagagcacaaccccttaagaggttgccagggccccagctgagcaacccatgctcctc 240  
Db 181 tcaagagagcacaaccccttaagaggttgccagggccccagctgagcaacccatgctcctc 240  
QY 181 tcaagagagcacaaccccttaagaggttgccagggccccagctgagcaacccatgctcctc 240  
Db 181 tcaagagagcacaaccccttaagaggttgccagggccccagctgagcaacccatgctcctc 240  
QY 241 gccacacagcctgtcgtcccttggttcctcgtggtcccccagatgctcgtgtgagacagc 300  
Db 241 gccacacagcctgtcgtcccttggttcctcgtggtcccccagatgctcgtgtgagacagc 300  
QY 241 gccacacagcctgtcgtcccttggttcctcgtggtcccccagatgctcgtgtgagacagc 300  
Db 241 gccacacagcctgtcgtcccttggttcctcgtggtcccccagatgctcgtgtgagacagc 300  
QY 301 cagtgagcctcagctggtccacacactctccggccccctgaaagtggcactgacagcagacag 360  
Db 301 cagtgagcctcagctggtccacacactctcttccggccccctgaaagtggcactgacagcagacag 360  
QY 301 cagtgagcctcagctggtccacacactctcttccggccccctgaaagtggcactgacagcagacag 360  
Db 301 cagtgagcctcagctggtccacacactctcttccggccccctgaaagtggcactgacagcagacag 360  
QY 361 ctccctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 420  
Db 361 ctccctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 420  
QY 361 ctccctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 420  
Db 361 ctccctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 420  
QY 421 cagagc 480  
Db 421 cagagc 480  
QY 421 cagagc 480  
Db 421 cagagc 480  
QY 481 ggtgtcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 540  
Db 481 ggtgtcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 540  
QY 481 ggtgtcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 540  
Db 481 ggtgtcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 540  
QY 541 ctgagagtcaccagtcaccagtcaccagtcaccagtcaccagtcaccagtcaccagtcaccagtcacc 600  
Db 541 ctgagagtcaccagtcaccagtcaccagtcaccagtcaccagtcaccagtcaccagtcaccagtcacc 600  
QY 541 ctgagagtcaccagtcaccagtcaccagtcaccagtcaccagtcaccagtcaccagtcaccagtcacc 600  
Db 541 ctgagagtcaccagtcaccagtcaccagtcaccagtcaccagtcaccagtcaccagtcaccagtcacc 600  
QY 601 ctgtgacatgtctgtacccctgagacagcagcgtggcagccaagggccctggggccacagcag 660  
Db 601 ctgtgacatgtctgtacccctgagacagcagcgtggcagccaagggccctggggccacagcag 660

Db 601 ctttgacatgctgtraccctgagacacacactgagccagccaccccttgaggccacgacg 660  
QY 661 tcggagagagcagcaccctgagagagcctgagcagctgcccggatctgacagcagccacagc 720  
Db 661 tcggagagagcagcaccctgagagagcctgagcagctgcccggatctgacagcagccacagc 720  
QY 721 gggcagcctggaactgtgtgcccggcggcgtgacacttggggagacactgcttggagcaggt 780  
Db 721 gggcagcctggaactgtgtgcccggcggcgtgacacttggggagacactgcttggagcaggt 780  
QY 721 gggcagcctggaactgtgtgcccggcggcgtgacacttggggagacactgcttggagcaggt 780  
Db 721 gggcagcctggaactgtgtgcccggcggcgtgacacttggggagacactgcttggagcaggt 780  
QY 781 gcaatcctggtgtgtggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 840  
Db 781 gcaatcctggtgtgtggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 840  
QY 781 gcaatcctggtgtgtggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 840  
Db 781 gcaatcctggtgtgtggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 840  
QY 841 catcaccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 900  
Db 841 catcaccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 900  
QY 841 catcaccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 900  
Db 841 catcaccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 900  
QY 901 agagcacaac 960  
Db 901 agagcacaac 960  
QY 901 agagcacaac 960  
Db 901 agagcacaac 960  
QY 961 gctgtcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1020  
Db 961 gctgtcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1020  
QY 961 gctgtcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1020  
Db 961 gctgtcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1020  
QY 1021 gccagccac 1080  
Db 1021 gccagccac 1080  
QY 1021 gccagccac 1080  
Db 1021 gccagccac 1080  
QY 1081 ggggattcactcactgtgtcctcagcagcagcagcagcagcagcagcagcagcagcagcagc 1140  
Db 1081 ggggattcactcactgtgtcctcagcagcagcagcagcagcagcagcagcagcagcagcagc 1140  
QY 1141 atcattcctcggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1200  
Db 1141 atcattcctcggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1200  
QY 1141 atcattcctcggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1200  
Db 1141 atcattcctcggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1200  
QY 1201 ccaagcctatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1260  
Db 1201 ccaagcctatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1260  
QY 1201 ccaagcctatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1260  
Db 1201 ccaagcctatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1260  
QY 1261 ggaactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1320  
Db 1261 ggaactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1320  
QY 1261 ggaactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1320  
Db 1261 ggaactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1320  
QY 1321 ggaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1380  
Db 1321 ggaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1380  
QY 1321 ggaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1380  
Db 1321 ggaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1380  
QY 1381 catctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1440  
Db 1381 catctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1440  
QY 1381 catctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1440  
Db 1381 catctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1440  
QY 1441 tgaacacccgcccctcaggggccccttctcctcctcctcctcctcctcctcctcctcctcctc 1500  
Db 1441 tgaacacccgcccctcaggggccccttctcctcctcctcctcctcctcctcctcctcctc 1500  
QY 1441 tgaacacccgcccctcaggggccccttctcctcctcctcctcctcctcctcctcctcctc 1500  
Db 1441 tgaacacccgcccctcaggggccccttctcctcctcctcctcctcctcctcctcctcctc 1500  
QY 1501 ggggaaacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1560  
Db 1501 ggggaaacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1560  
QY 1501 ggggaaacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1560  
Db 1501 ggggaaacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1560  
QY 1561 gcaac 1620  
Db 1561 gcaac 1620  
QY 1561 gcaac 1620  
Db 1561 gcaac 1620  
QY 1621 gttgttctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1680  
Db 1621 gttgttctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1680  
QY 1621 gttgttctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1680  
Db 1621 gttgttctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1680  
QY 1681 caaacacccctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1740  
Db 1681 caaacacccctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1740

Query Match 2.0%; Score 38; DB 12; Length 1704;

Best Local Similarity	100.0%	Pred. No. 4	5e-10	
Matches	38:	Conservative	0:	Mismatches 0: Indels 0: Gaps 0:
Qy	1025	gcccacctggacatctggaaagtcacagcgccctggatgaa	1062	
Db	941	gcccaactggacatctggaagtcacagcgccctggatgaa	978	
RESULT	3			
AF075706	848 bp	mRNA	VRT	19-DEC-1998
LOCUS				
DEFINITION	Gallus gallus ets domain protein (ERM) mRNA, partial cds.			
ACCESSION	AF075706			
NID	93869357			
VERSION	AF075706.1	GI:3869357		
KEYWORDS				
SOURCE	Chicken.			
ORGANISM	Gallus gallus			
REFERENCE	Eukariota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.			
AUTHORS	1 (bases 1 to 848) Lin,J.H., Saito,T., Anderson,D.J., Lance-Jones,C., Jessell,T.M. and Arber,S.			
TITLE	Functionally related motor neuron pool and muscle sensory afferent subtypes defined by coordinate ERS gene expression			
JOURNAL	Cell 95 (3), 393-407 (1998)			
MEDLINE	99029846			
REFERENCE	2 (bases 1 to 848)			
AUTHORS	Lin,J.H., Saito,T., Anderson,D.J., Lance-Jones,C., Jessell,T.M. and Arber,S.			
TITLE	Direct Submission			
JOURNAL	Submitted (01-JUL-1998) Center for Neurobiology & Behavior, Columbia University, 701 W. 168th St. HSC 1013, New York, NY 10032, USA			
FEATURES				
source	Location/Qualifiers			
gene	1..848			
CDS	/organism="Gallus gallus" /db_xref="taxon:9031" <1..848 /gene="ERM" <1..712 /gene="ERM" /codon_start=2 /product="ets domain protein" /protein_id="AAC97201.1" /db_xref="PID:g3869358" /db_xref="GI:3869358" /translation="RPPAPRPRRGSPQPMGIRKQEPDPCIDSEVPNCSSVYVGGLF PSSNDGYSYEDTRLYFDIDTVYVERLEKRYKQPTLYRSPPTQRRGSLQMQFLYT LDDPANRHFIAWTGRGHEFLIEPEEYARWGIQKRPANIKDLSRLYIEKTI MOKAIGERYVYKFEVCDPALEFSMAVPDNQRFLAEPDCHVSEEDTLPLTHFEDSPAY LLEVEHCSSLVYAGFPAY"			
BASE COUNT	192 a 265 c 245 g 146 t			
ORIGIN				
Query Match	1.6%; Score 30; DB 4; Length 848;			
Best Local Similarity	100.0%; Pred. No. 1			
Matches	30: Conservative 0: Mismatches 0: Indels 0: Gaps 0:			
Qy	1310	gccatgacctagcagacagctgagcgctcc	1339	
Db	425	gccatgacctagcagacagctgagcgctcc	454	
RESULT	4			
AF075708	2508 bp	mRNA	VRT	19-DEC-1998
LOCUS				
DEFINITION	Gallus gallus ets domain protein (PEA3) mRNA, complete cds.			
ACCESSION	AF075708			
NID	94033764			
VERSION	AF075708.1	GI:4033764		
KEYWORDS				

SOURCE Chicken.  
ORGANISM Gallus gallus  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.  
AUTHORS Lin, J.H., Salto, T., Anderson, D.J., Lance-Jones, C., Jessell, T.M. and Arber, S.  
TITLE Functionally related motor neuron pool and muscle sensory afferent subtypes defined by coordinate ETS gene expression  
JOURNAL Cell 95 (3), 393-407 (1998)  
MEDLINE 99029846  
REFERENCE 2 (bases 1 to 2508)  
AUTHORS Lin, J.H., Salto, T., Anderson, D.J., Lance-Jones, C., Jessell, T.M. and Arber, S.  
TITLE Direct Submission  
JOURNAL Submitted (01-JUN-1998) Center for Neurobiology & Behavior, Columbia University, 701 W. 168th St. HSC 1013, New York, NY 10032, USA  
COMMENT On Dec 19, 1998 this sequence version replaced g1.3869361.  
FEATURES  
source location/Qualifiers  
1..2508  
/organism="Gallus gallus"  
/db\_xref="taxon:9031"  
1..2508  
/gene="PEA3"  
1..1011  
/gene="PEA3"  
/codon\_start=1  
/product="ets domain protein"  
/protein\_id="AAC97202.1"  
/db\_xref="PDB:4033765"  
/db\_xref="GI:4033765"  
/translation="VHNNHRIVLTQLYLFDSPPAYDOPRPYIKSPDPTPIOSPOOF  
PPOORNEFLPSPGPHPTSMGYLIENSSAOAPVEMCHSSSGMADAPYORPM  
SEOCLOCPPOGKOFHPDPOYEOAGSCQAPAYIKOEDVDYDSCGCTYDQMRPF  
ADYCVVPEKREPDIKOEVGRGSPYORRGSLQWLFVALDDPTNSHTAWGR  
GMEFKLIEPEVARLKGIOKRPAMNTDKLSRLTYEYGIQKVAQERYVYKFC  
PEALFSLAFPDNORPALKALELRQISEDDVPLSLDESTAYVLPDLSLPOLYKGY  
TY"  
BASE COUNT 563 a 715 c 672 g 557 t 1 others  
ORIGIN  
Query Match 1.58; Score 29; DB 4; Length 2508;  
Best Local Similarity 100.0%; Pred. No. 4.8e-05;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1310 gccatgaactacgaacactgacgcgtc 1338  
Db 724 gccatgaactacgaacactgacgcgtc 752  
RESULT 5  
AC007818/c AC007818 . 60056 bp DNA HMG 14-JUN-1999  
LOCUS Drosophila melanogaster chromosome 3 clone BACR02M05 (D785) RCI-98  
DEFINITION 02.M.5 map 98B-98B strain Y; cn bw sp, WORKING DRAFT SEQUENCE, 20  
unordered pieces.  
AC007818  
AC007818  
95053146  
AC007818.1 GI:5053146  
KEYWORDS HMG; HMG5-PHASE1.  
SOURCE Fruit fly.  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE 1 (bases 1 to 60056)  
AUTHORS Ceiniker, S.E., Agbayani, A., Arcaina, T.T., Baxter, E., Blazef, R.G.,  
Butenhoff, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,  
Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,  
Hoskins, R.A., Houston, K.A., Hummasti, S.R., Katta, K., Kearney, L.,  
Kim, E., Lee, B., Lewis, S., Li, P., Lomocan, M.A., Mazda, P.,  
K.M., E., Lee, B., Lewis, S., Li, P., Lomocan, M.A., Mazda, P.,

Moshrefi, A.R., Moshrefi, M., Nixon, K., Pacleb, J.M., Park, S.,  
Pfeiffer, B., Poon, L., Sequelra, A., Sethi, H., Snir, E.,  
Svirkas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zieran, L.L. and  
Rubin, G.M.  
TITLE Sequencing of Drosophila melanogaster  
JOURNAL Unpublished  
AUTHORS 2 (bases 1 to 60056)  
REFERENCE Ceiniker, S.E., Agbayani, A., Arcaina, T.T., Baxter, E., Blazef, R.G.,  
Butenhoff, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,  
Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,  
Hoskins, R.A., Houston, K.A., Hummasti, S.R., Katta, K., Kearney, L.,  
Kim, E., Lee, B., Lewis, S., Li, P., Lomocan, M.A., Mazda, P.,  
Moshrefi, A.R., Moshrefi, M., Nixon, K., Pacleb, J.M., Park, S.,  
Pfeiffer, B., Poon, L., Sequelra, A., Sethi, H., Snir, E.,  
Svirkas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zieran, L.L. and  
Rubin, G.M.  
TITLE Direct Submission  
JOURNAL Submitted (14-JUN-1999) Drosophila Genome Center, Lawrence Berkeley  
Laboratory, MS 64-121, Berkeley, CA 94720, USA  
COMMENT For further information about this sequence, including its location  
and relationship to other sequences, please visit our sequence  
archive web site (<http://www.fruitfly.org/sequence/>) or send email  
to [bdgs@fruitfly.berkeley.edu](mailto:bdgs@fruitfly.berkeley.edu). All contigs in this submission meet  
the following cutoffs: length >= 400 bases, phrap computed error  
rate <= 1/10.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 20 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1  
893: contig of 893 bp in length  
894  
973: gap of unknown length  
974  
1964: contig of 991 bp in length  
1965  
2044: gap of unknown length  
2045  
3652: contig of 1608 bp in length  
3653  
3732: gap of unknown length  
3733  
3772: contig of 1540 bp in length  
3773  
3532: gap of unknown length  
3533  
6476: contig of 1124 bp in length  
6477  
6556: gap of unknown length  
6557  
8659: contig of 2113 bp in length  
8660  
8749: gap of unknown length  
8750  
11281: contig of 2532 bp in length  
11282  
11361: gap of unknown length  
11362  
13375: contig of 1914 bp in length  
13376  
13355: gap of unknown length  
13356  
16154: contig of 2799 bp in length  
16155  
16234: gap of unknown length  
16235  
18409: contig of 2175 bp in length  
18410  
18489: gap of unknown length  
18490  
20047: contig of 1558 bp in length  
20048  
20127: gap of unknown length  
20128  
21987: contig of 1860 bp in length  
21988  
22067: gap of unknown length  
22068  
24814: contig of 2847 bp in length  
24815  
24954: gap of unknown length  
24915  
28786: contig of 3792 bp in length  
28787  
28866: gap of unknown length  
28867  
32920: contig of 4054 bp in length  
32921  
33000: gap of unknown length  
33001  
36522: contig of 3322 bp in length  
36523  
36602: gap of unknown length  
36603  
41889: contig of 5287 bp in length  
41890  
41969: gap of unknown length  
41970  
46511: contig of 4542 bp in length  
46512  
46591: gap of unknown length  
46592  
51474: contig of 4883 bp in length  
51475  
51554: gap of unknown length  
51555  
60056: contig of 8502 bp in length.  
Location/Qualifiers

FEATURES

source 1. 60056  
/organism="Drosophila melanogaster"  
/strain="y: cn bw ap"  
/db\_xref="taxon:7227"  
/clone="BACR02M05 (D785) RPCI-98 02.M.5"  
/chromosome="3"  
/clone\_lib="RPCI-98 (Roswell Park Cancer Institute  
Drosophila melanogaster BAC library, partial EcORI in  
PBAC3.6"  
/map="98B-98B"  
BASE COUNT 1556 a 1359 c 1365 g 15945 t 1523 others  
ORIGIN

Query Match 1.5%: Score 28; DB 35; Length 60056;  
Best Local Similarity 100.0%; Pred. No. 0.00018;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1159 catccacctgtgacgttcctcaaggag 1186  
|||||  
DB 54075 CATCCACCTGTGCACTCTCTCAAGAG 54048

RESULT 6  
DRODERS4A 362 bp DNA INV 05-AUG-1992  
LOCUS Drosophila melanogaster D-ets-4 DNA binding domain protein gene, 3'  
DEFINITION end  
ACCESSION M88474  
M88474  
NID 9157195  
VERSION M88474.1 GI:157195  
KEYWORDS DNA-binding domain protein.  
SOURCE Drosophila melanogaster DNA.  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE 1 (bases 1 to 362)  
AUTHORS Chen, T., Bunting, M., Karim, F.D. and Thumel, C.S.  
TITLE Isolation and characterization of five Drosophila genes that encode  
an ets-related DNA binding domain  
JOURNAL Dev. Biol. 151, 176-191 (1992)  
MEDLINE 92249640  
FEATURES  
source Location/Qualifiers  
1..362  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/map="Chromosome 3R 98B"  
1..345  
/gene="Ets98B"  
/note="Ets at 98B"  
/allele=""  
/db\_xref="FlyBase:FBgn0005659"  
1..345  
/partial  
/gene="Ets98B"  
/codon\_start=1  
/product="D-ets-4 DNA binding domain protein"  
/protein\_id="AA28451.1"  
/db\_xref="PIR:G157196"  
/db\_xref="PIR:G157196"  
/db\_xref="GI:157195"  
/translation="TNAAGGATATKRPNGRTGGGGSHILWFLKELASPOVNGT  
AIRWIDRSKGIKFIKEDSVVAKLMGRNRKRAMYDKLSRSIROYKKGIMKTKERSO  
RLVYOFCHPYGSO"  
BASE COUNT 90 a 99 c 113 g 60 t  
ORIGIN Chromosome 3R, 98B.

Query Match 1.5%: Score 28; DB 36; Length 362;  
Best Local Similarity 100.0%; Pred. No. 0.00017;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1159 catccacctgtgacgttcctcaaggag 1186

DB 75 CATCCACCTGTGCACTCTCTCAAGAG 102  
|||||

RESULT 7  
HSDNAFEV3 1403 bp DNA PRI 23-APR-1997  
LOCUS H.sapiens FEV gene, exon 3.  
DEFINITION Y08979  
ACCESSION Y08979  
NID 91841699  
VERSION Y08979.1 GI:1841699  
KEYWORDS ets gene family.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 1403)  
AUTHORS Peter, M., Couturier, J., Pacquement, H., Michon, J., Thomas, G.,  
Magdelenat, H. and Delattre, O.  
TITLE A new member of the ETS family fused to EWS in Ewing tumors  
JOURNAL Oncogene 14 (10), 1159-1164 (1997)  
MEDLINE 97236556  
REFERENCE 2 (bases 1 to 1403)  
AUTHORS Delattre, O.  
TITLE Direct Submission  
JOURNAL Submitted (10-OCT-1996) O. Delattre, Institut Curie, Inserm U434,  
26 Rue Duhm, Paris 75231 Cedex 05, Paris 75231 Cedex 05, FRANCE  
REMARK revised by [3]  
AUTHORS Delattre, O.  
TITLE Direct Submission  
JOURNAL Submitted (13-FEB-1997) O. Delattre, Institut Curie, Inserm U434,  
26 Rue Duhm, Paris 75231 Cedex 05, Paris 75231 Cedex 05, FRANCE  
COMMENT On Feb 15, 1997 this sequence version replaced gi:1834485.  
FEATURES  
source Location/Qualifiers  
1..1403  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/germline  
1..240  
/number=2  
/evidence=experimental  
gene 241..1403  
/gene="FEV"  
241..1403  
/gene="FEV"  
/note="ETS family member"  
/number=3  
/evidence=experimental  
BASE COUNT 210 a 551 c 376 g 266 t  
ORIGIN

Query Match 1.3%: Score 25; DB 10; Length 1403;  
Best Local Similarity 100.0%; Pred. No. 0.0083;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1312 catgacctgacacagctgagccgc 1336  
|||||  
DB 398 CATGAACCTACGACAGCTGAGCCGC 422

RESULT 8  
HSRNAFEV 1901 bp mRNA PRI 23-APR-1997  
LOCUS H.sapiens mRNA for FEV protein.  
DEFINITION Y08976  
ACCESSION Y08976  
NID 91834508  
VERSION Y08976.1 GI:1834508  
KEYWORDS ets gene family.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;

```

REFERENCE 1 (bases 1 to 1901)
AUTHORS Peter,M., Couturier,J., Paquement,H., Michon,J., Thomas,G.,
          Magdelenat,H. and Delattre,O.
TITLE A new member of the ETS family fused to EWS in Ewing tumors
JOURNAL Oncogene 14 (10), 1159-1164 (1997)
MEDLINE 9723656
REFERENCE 2 (bases 1 to 1901)
AUTHORS Delattre,O.
TITLE Direct Submission
JOURNAL Submitted (10-OCT-1996) O. Delattre, Institut Curie, Inserm U434,
          26 Rue Duhm, Paris 75231 Cedex 05, Paris 75231 Cedex 05, FRANCE
FEATURES
source
1. 1901
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/cell_line="ewing"
585..1301
/function="DNA binding protein"
/notes="ETS family member"
/codon_start=1
/product="E7V protein"
/protein_id="CAA70169.1"
/db_xref="PID:e291921"
/db_xref="PID:g1834509"
/db_xref="GI:1834509"
/db_xref="SPTREMBL:O99581"
/translation="NR0SGASORPLINNYLDPVGGDLFKDKGNPSMGPLSPAVOKSS
GOIQLOMFLLEADRLANAGCIAMGSGGERLDPDEVARWGERKSPMMYDKLS
RALRIYDKNINISVHKRYAVREDPGLAQCOPPAHAAAAAADAADGAL
KPLPGGLAPLPFPGLSKLNLMASAGVAPAGFSYWPGPAPATAAATAALYPSISQ
PPGPGFAGVAAASHLGGHYH"
711..962
/notes="DNA binding domain"
BASE COUNT 302 a 707 c 575 g 317 t
ORIGIN
misc_feature
Query Match 1.3%; Score 25; DB 10; Length 1901;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 25: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1312 catgaactagacaagctgagccg 1336
|||||
Db 869 CATGAACACTAGACAGCTGAGCCG 893
|||||
RESULT 9
DRODETS6A 480 bp DNA INV 05-AUG-1992
LOCUS Drosophila melanogaster D-ets-6 DNA binding domain protein gene,
          partial cds.
ACCESSION M88475
M88475
157197
VERSION M88475.1 GI:157197
KEYWORDS DNA-binding domain protein.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
          Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
          Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 480)
AUTHORS Chen,T., Buntling,M., Karim,F.D. and Thummel,C.S.
TITLE Isolation and characterization of five Drosophila genes that encode
          an ets-related DNA binding domain
JOURNAL Dev. Biol. 151, 176-191 (1992)
MEDLINE 92249640
FEATURES
source
1. 480
/location/Qualifiers
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/map="Chromosome 2L 21C."

```

```

Intron
1. 95
/gene="D-ets-6 DNA-binding domain"
96..288
/gene="Ets21C"
/db_xref="FlyBase:FBgn0005660"
join(96..288,366..480)
/gene="Ets21C"
/notes="Ets at 21C"
/allele=""
/db_xref="FlyBase:FBgn0005660"
join(96..288,366..480)
/partial
/gene="Ets21C"
/codon_start=3
/product="D-ets-6 DNA binding domain protein"
/protein_id="AAA28452.1"
/db_xref="PID:g157198"
/db_xref="GI:157198"
/translation="SGGQIQLOMFLLEADSSNANAISWEGQSGEFLIDPEVARR
WGRRAKPNMNYDKLSRALRLRYDDKNIMTKVGRYAVKFDHGLAACOAQAGGDP"
289..365
/gene="D-ets-6 DNA-binding domain"
366..480
/gene="Ets21C"
/db_xref="FlyBase:FBgn0005660"
Chromosome 2L, 21C.
BASE COUNT 101 a 138 c 142 g 99 t
ORIGIN
Chromosome 2L, 21C.
Intron
289..365
/gene="D-ets-6 DNA-binding domain"
366..480
/gene="Ets21C"
/db_xref="FlyBase:FBgn0005660"
99 t
Chromosome 2L, 21C.
Query Match 1.3%; Score 24; DB 36; Length 480;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 24: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1312 catgaactagacaagctgagccg 1335
|||||
Db 256 CATGAACACTAGACAGCTGAGCCG 279
|||||
RESULT 10
AC004573/c 85095 bp DNA INV 30-JAN-1999
LOCUS Drosophila melanogaster. chromosome 2L, region 21C5-21D1, pl clone
          DS07610, complete sequence.
ACCESSION AC004573
94204255
VERSION AC004573.1 GI:4204255
KEYWORDS fruit fly.
SOURCE Drosophila melanogaster
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
          Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
          Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 85095)
AUTHORS Celniker,S.E., Adayani,A., Arcana,T.T., Baxter,E., Blazej,R.G.,
          Butenhoif,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
          Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
          Hoskins,R.A., Houston,K.A., Hummstil,S.R., Karra,K., Kearney,L.,
          Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
          Mosherel,A.R., Mosherel,M., Nixon,K., Pacled,J.M., Park,S.,
          Pfeiffer,B., Poon,L., Sequerra,A., Sethi,H., Shit,E.,
          Svitskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
          Rubin,G.M.
TITLE Sequencing of Drosophila chromosome 2L, region 21C5-21D1
JOURNAL Unpublished (1998)
REFERENCE 2 (bases 1 to 85095)
AUTHORS Celniker,S.E., George,R.A., Galle,R., Svitskas,R.R., Hoskins,R.A.,
          Agdeyani,A., Arcana,T.T., Baxter,E., Blazej,R.G., Chavez,C.,
          Chew,M., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Houston,K.A.,
          Hummstil,S.R., Karra,K., Kearney,L., Kim,S.H., Lee,B.,
          Lomotan,M.A., Mak,J., Mazda,P., Mok,M.S., Mosherel,A.R.,
          Mosherel,M., Nixon,K., Pacled,J.M., Park,S., Pfeiffer,B.,
          Shit,E., Twomey,B., Wan,K.H., Whiteclaw,K.R., Yee,A., Zhang,R.,

```



```

misc_feature /note="MIR repeat: matches 146..76 of consensus"  

repeat_region /note="match: GSS A006992"  

repeat_region /note="MULT2D repeat: matches 553..1 of consensus"  

repeat_region /note="MIR repeat: matches 20..141 of consensus"  

misc_feature /note="match: GSS B90081"  

repeat_region /note="MER28 repeat: matches 1..434 of consensus"  

repeat_region /note="THEIC repeat: matches 1..371 of consensus"  

repeat_region /note="MIR2 repeat: matches 134..70 of consensus"  

repeat_region /note="MULT2G repeat: matches 388..324 of consensus"  

repeat_region /note="MULT2D repeat: matches 553..505 of consensus"  

Query Match 1.28; Score 22; DB 9; Length 115835;  

Best Local Similarity 100.0%; Pred.No.0.4;  

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0  

Oy 1525 ctgcctgacccctccagagccc 1546  

|||||  

Db 88123 CTGCTGTGACCTTCAGAGGCC 88102  

  

RESULT 12  

LOCUS MNNETRN 2658 bp mRNA ROD 13-MAR-1995  

DEFINITION M.musculus net mRNA.  

ACCESSION Z32815  

MID 9479112  

VERSION 232815.1 GI:479112  

KEYWORDS Net; ras gene.  

SOURCE house mouse.  

ORGANISM Mus musculus  

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  

Rodentia; Sciurognathi; Muridae; Murinae; Mus.  

REFERENCE 1 (bases 1 to 2658)  

AUTHORS Giovane,A., Pintas,A., Maiza,S.M., Sobieszczuk,P. and Waslyk,B.  

TITLE Net, a negative factor switch to positive by Ras  

JOURNAL Unpublished  

REFERENCE 2 (bases 1 to 2658)  

AUTHORS Giovane,A.  

TITLE Direct Submission  

JOURNAL Submitted (29-APR-1994) Antoine Giovane,  

CNRS-IGME, INSERM-U.184, Institut de Chimie, Biologique, 11 rue  

Humann, Strasbourg, 67085 Strasb. Cedex, France  

3 (bases 1 to 2658)  

AUTHORS Giovane,A., Pintas,A., Maiza,S.M., Sobieszczuk,P. and Waslyk,B.  

TITLE Net, a new ets transcription factor that is activated by Ras  

JOURNAL Genes Dev. 8 (13), 1502-1513 (1994)  

MEDLINE 95047310  

FEATURES  

source location/qualifiers  

1..2658  

/organism="Mus musculus"  

/db_xref="taxon:10090"  

/dev_stage="10 days"  

/tissue_type="embryo"  

/clone_lib="lambda zap2"  

295..1524  

/codon_start=1  

/product="Net"  

/protein_id="CAA83676.1"  

/db_xref="PID:g479113"  

/db_xref="GI:479113"  

/db_xref="MD:MGI:101762"  

/db_xref="SWISS-PROT:P41971"  

/translation="MESATITWQFLHLHLLDKKHBNLICTWSNDGEFKLLAAEEVAKI"

```



```

ESLLLOOGDCVSPDEGRVHGHGSLSSKASRNRYLLSGVLSSTINSLENAAPAFK
IKTKPLKEPDOSPVEEVRVIRFVTKNKDKHTRPVMSLPSTSEAAAAAAPLAK
SVSAKISLSLMPNAAVSYSVPSSSRSGSLSPDPSLPSRHSLSLEACHESDSELP
NLSGSKRTKSPSLPPKRRKPGLEISAPQLLSTDIDGSLNLSNPALPSGGLTAAFT
AOTSGGLFASPSPSLPIHFWSSLSPPVAPLSPARLOGNTLFOEPTLNGHMPVLP
LDRAPSVLSPSSONS"

```

Query Match	1.2%	Score 22;	DB 12;	Length 2658;
Best Local Similarity	100.0%	Pred. No. 0.39;		
Matches 22: Conservative	0;	Mismatches	0;	Indels 0;
				Gaps 0;

Qy	1312	catgaactacgacaaagctgagc	1333
Db	453	CATGAAGTACGACAAAGCTGAGC	474

RESULT	13					
MUSTRANONC		2211 bp	mRNA			19-JUL-1994
LOCUS						
DEFINITION		Mus musculus ERP mRNA, complete cds.				
ACCESSION		L19953				
NID		9436184				
VERSION		L19953.1				
KEYWORDS		transcription factor.				
SOURCE		Mus musculus (Strain BALB/c, sub-species domesticus) (library				

ORGANISM	REFERENCE AUTHORS TITLE	JOURNAL MEDLINE FEATURES
<i>Mus musculus</i>	Eukaryote, Metazoa, Chordata, Craniata, Vertebrata, Mammalia, Euthalia, Rodentia, Sciurognathi, Muridae, Murinae, Mus.	
1 (bases 1 to 2211)	Lopez, M., Oettingen, P., Akbarali, Y., Dendorfer, U. and Libermann, T.A	
	ERP, a new member of the ets transcription factor/oncoprotein family: cloning, characterization, and differential expression during B-lymphocyte development	
	Mol. Cell. Biol. 14 (5), 3292-3309 (1994)	
94217726	Location/Qualifier	

```

/organism="Mus musculus"
/strain="BALB/c"
/sub_species="domesticus"
/db_xref="taxon:10090"
/sex="male"
/tissue_type="lung"
/cisue_lib="lambda-gtl of Clontech"
302. .2109
/gene="rrp"
302. .1531
CDS

```

```

/codon_start=1
/protein_id="AA19872.1"
/db_xref="PID:9436185"
/db_xref="GI:436185"
/translation="MESAITLMQFLHLLLDKQHEHLICMTSNDGEFKLKAEEVAKI
WGLKKNKTNMNTDKLSRALRYTYDKNTIKATGCFATYKTSFPIKLMQDHAPEIKS
ESLLLDQDCCKVSPGEGREVAHRHGJSLKSAARNETLHSGLSFTFLNSLQNAPEKFK
IKTKLEPECDSPVEVEATVIRVETKTKTRKHIRPWSLPTSETIAAASAFLEA
SVSAKISLIMLPNNAASVSASPSGSRSLSPSLDPSEHSLFTEACCHSDSDFSL
NLSGSKTSPSLPQKGGKPGELISAQOLLSCGDISLINSAPLNGSGSTVPAFF
AQTSGCLFIASSPLLPSTHFWSLSLPAVLPSPARLQGNITLQFPTLLNGHMPVLP
LDRAAPSPVLPSSGKRS"

```

```
misc_feature      320..553
                  /gene="ERP"
                  /note="ETS domain"
polyA_signal      2195..2199
                  /gene="ERP"
BASE COUNT        560 a      504 g      534 t
ORIGIN
```

	Query Match	1.2%; Score 22;	DB 12;	Length 2211;
	Best Local Similarity	100.0%;	Pred. No. 0.97;	
Matches	22:	Conservative	0;	Mismatches
			0;	Indels
			0;	Gaps
Oy	1312 catgaactacgacaagctgagc	1333		
D8	460 CATGAACACTACGACAAAGCTGAGC	481		

RESULT	14
LOCUS	S82864
DEFINITION	S82864 1526 bp mRNA ERNA
ACCESSION	nt]. Etk-3-Ets transcription factor [mice, 16-day embryos, mRNA, 1526 nt].
NID	S82864
VERSION	G1836129
KEYWORDS	S82864.1 GI:1836129
SOURCE	Mus sp. 16-day embryos.
ORGANISM	Mus sp.

REFERENCE ARTHOSS TITLE	JOURNAL MEDLINE REMARK
1 (bases 1 to 1526) Nozaki, M., Onishi, Y., Kanno, N., Ono, Y. and Fujimura, Y. Molecular cloning of Elk-3, a new member of the Ets family expressed during mouse embryogenesis and analysis of its transcriptional repression activity DNA Cell Biol. 15 (10), 855-862 (1996)	Genbank staff at the National Library of Medicine created this entry [NCBI gblbseq.179297] from the original journal article. This sequence comes from Fig. 1.

FEATURES	Location/Qualifiers
source	1..1526
	/organism="Mus sp."
	/db_xref="taxon:10095"
gene	94..1323
	/gene="Elk-3"
CDS	94..1323

```

/note="Ets transcription factor; This sequence comes from
Fig. 1"
/codon_start=1
/product="E1x-3"
/protein_id="AAB46893.1"
/db_xref="PID:G1836130"
/db_xref="GI:1836130"
/translation="MESAITLWOLFLLHLDQKHEHLICWTSNDGEFKLLAAEEVAKI
WGLAKNTNNMYDKLSRLRYVYDKNITIKYIGQGFYVSPDILKMDPAVVISL
ESLLDQDGCYVSPGEGREVRHKGSLSKASRNEYLNHGILSFTINLQNAPEFKR
IKTKLEPCDDSPVVEEVRTVIRFVNTKDKHITRPVSLPSTLEATAAASDLEP
SVKALISLIMLPNAAVSASVSPSSRSLSLSPDPSLPENHSLFLEAACHSASL
NQTSGSTKSPSLPPKCKKPGGLEISAPOLLGCDGDISLINSALPDSGSLTPFF
AOTSGEFLASSPLPSIHFWSSLSVPAPLSPARLQGNTLQFQFTLLNGHMPVLP
LDRAPSVILSPSSQKS"

```

Query Match	1.2%	Score 22:	DB 12:	Length 1526:
Best Local Similarity	100.0%	Pred. No.	0.39:	
Matches 22:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0

```

Oy 1312 catgaactacgacaaagctgagc 1333
      |||||
Db 252 CATGAAGTACGACAAAGCTGAGC 273

```

RESULT	15	
LOCUS	PFMAL3P4/c	
DEFINITION	PFMAL3P4	113899 bp DNA
ACCESSION	AL008970	INV 29-APR-1999
		Plasmodium falciparum MAL3P4, complete sequence.

```

NID          94493924
VERSION      AL008970.2
KEYWORDS     GI:4493924
SOURCE       malaria parasite P. falciparum.
ORGANISM     Plasmodium falciparum
REFERENCE    Eukaryota; Alveolata; Apicomplexa; Hemosporida; Plasmodium.
AUTHORS      1 (bases 1 to 113899)
              Hamlin, N., Bowman, S., Churcher, C., Lawson, D., Quall, M. and
              Barrell, B.
TITLE        Direct Submission
JOURNAL      Submitted (15-APR-1999) P.falciparum Genome Sequencing Consortium,
              The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
              CB10 1SA, UK
COMMENT      On Mar 24, 1999 this sequence version replaced gi:3763998.
              For more information about this sequence or the Malaria Project,
              see http://www.sanger.ac.uk/Projects/P_falciparum. IMPORTANT: This
              sequence is not the entire insert of clone MAL3P4. It may be
              shorter because we only sequence overlapping sections once, or
              longer because we arrange for a small overlap between neighbouring
              submissions.

FEATURES
    source          1. 113899
                    /organism="Plasmodium falciparum"
                    /strain="3D7"
                    /db_xref="taxon:5833"
                    /chromosome="3"
                    /clone="MAL3P4"
                    /size="6248"
                    /gene="MAL3P4.26"
                    /note="MAL3P4.26"
                    /note="predicted using hexexon: MAL3P4.26 (PFC0435w),
                    Hypothetical protein, len: 1295 aa"
                    /codon_start=1
                    /protein_id="CA15616.1"
                    /db_xref="PID:e1332567"
                    /db_xref="PID:g3764023"
                    /db_xref="GI:3764023"
                    /db_xref="SPTREMBL:O77361"
                    /translation="MKVGIIFCLFEFFVVGACNNVERIFKRIKRTKFIILEPIV
                    DLSTSENLFTLLDLVDKRLYLTLDESILNLENNLSEIFLLVDTYKIKENEDN
                    KNIRYIFLGTSFRIHPLNEFYLRKLNKYINGNIYKGNVDIRGLIEYKNEIEEK
                    KLEOKLNTKIKNNNNNNNNNSKFSKDGNDENKNDLYNPSDKLNNNDIDVHE
                    LLEIIITKEKREFLDDDDSDNSKYTLKDEYKRTKGFITGIGFNDIPSVIHHNF
                    DKNTLSPSLNSGITLITLLKNITENSIILSNNEKQSHIDYIYEVKYLKENLAV
                    RLTHSEVNCLEEDNIHLNDPNEFEIYIYOVLALEKDYKNTIEKQYKIGHEHNV
                    RHEITSSGEENLRNTRKNNNNNNNNSEDAIELLSYFNVEYPISTCMYSIR
                    SKHESLMDYKHYHNLNDIKLHYKEEIEHFNSEIEYKMLRNKYKDTLLEE
                    HENLVYTKNLIGIKTINTEERIPKIKTYDNKENTQIIFTEYNDKIKENITFG
                    FYNSLLONALENDNDITLITMSDESKQYDLNYSKYTSKEGICELKHHITYY
                    EYVYMKSEKKYFIADDDPFVYKULIDVTNLTLTSCSKTYMDTKIKSIDVKE
                    NEALFLONPKRTKFLPLSYTKDTPAKTITQLTKDYVPYCOGILSKHKHNSDDDD
                    HDHVGKNNNDSTNHODIEKQVNIINNNNNNNNKAKSIPYIGRRSYTPTSE
                    NEFYDVLGTGAGAILNDEYAKRIEKECTCPTSSMDMIFGKWAELGILANF
                    EGYFQNSPLDNNKRYINTLVPITHYRLNKRRTKESRDYFNVLVYNNNDKONKI
                    YVYLDNRHKNMIDNVEFYVYNTDKKRYVTKTEHNAADNSKRNKRNQKUNTI
                    OGKNNVNDENVADENYKGDENYKGEYKKGDEYKGDENYKGDENYKGDENYK
                    NIKGDNNVNDENVADENYKGDENYKGEYKKGDEYKKGDEYKGDENYKGDENYK
                    KKSVPNNITDKNEDTTRYIKNENKTYINRMQESGKYQLFDINKFFKKEIGHYFQ
                    KIRKNEKAKKEKERNQOLKOKDYTNNTFHTSNMONGNMOOKMGVYQOENEDDF
                    DQPELEDAINPMDEYENENLSEFDEGEDEYDDYDVTINADKLINDQNK
                    HLYEQIKDIAQPVNFONDONSNTFDFIDEL"
                    complement(7161..13970)
                    /gene="MAL3P4.25"
                    complement(7161..13970)
                    /gene="MAL3P4.25"
                    /note="predicted using hexexon: MAL3P4.25 (PFC0440c),
                    Helicase, len: 2270 aa; similarity to helicases. C.elegans
                    5.9e-31; 34% in 199 aa overlap."
                    /codon_start=1
                    /protein_id="CA15615.1"
                    /db_xref="PID:e1332566"

gene
CDS
    /gene="MAL3P4.25"
    /note="predicted using hexexon: MAL3P4.25 (PFC0440c),
    Helicase, len: 2270 aa; similarity to helicases. C.elegans
    5.9e-31; 34% in 199 aa overlap."
    /codon_start=1
    /protein_id="CA15615.1"
    /db_xref="PID:e1332566"

gene
CDS
    /gene="MAL3P4.23"
    /note="predicted using hexexon: MAL3P4.24 (PFC0445w),
    Hypothetical protein, len: 111 aa"
    /codon_start=1
    /protein_id="CA15613.1"
    /db_xref="PID:e1332564"
    /db_xref="PID:g3764020"
    /db_xref="GI:3764020"
    /db_xref="SPTREMBL:O77358"
    /translation="MLHGISTISEKINVLSEYKRTINIOSLEKGIETIEGCGFRIQ
    CYDITGKIKIFVHKDDLNTLNTLKRVEYELSDIILNPFIDIMPIRSANVFQI
    EKLFSNTS"
    19545..19874
    /gene="MAL3P4.23"
    19545..19874
    /gene="MAL3P4.23"
    /note="predicted using hexexon: MAL3P4.24 (PFC0445w),
    Hypothetical protein, len: 110 aa"
    /codon_start=1
    /protein_id="CA15614.1"
    /db_xref="PID:e1332565"
    /db_xref="PID:g3764021"
    /db_xref="GI:3764021"
    /db_xref="SPTREMBL:O77359"
    /translation="MSMFLNIIILIDPAASVAFLLITFLMININSESLISQAHRENGK
    KALVAIILVIVLVLFYIKAVKNNKRLTYTFPKKRAPIYVOLASTYLSASDEYE
    OYELNKI"

```

gene  
CDS

20968..23328  
/gene="MAL3P4.22"  
20968..23328  
/gene="MAL3P4.22"  
/note="predicted using hexExon; MAL3P4.22 (PFC0455w),  
Hypothetical protein, len: 787 aa"  
/codon\_start=1  
/protein\_id="CAA15612.1"  
/db\_xref="PID:g1332563"  
/db\_xref="PID:g13764019"  
/db\_xref="GI:3764018"  
/db\_xref="SPTREMBL:O77357"

gene  
CDS

24170..25001  
/gene="MAL3P4.21"  
Join(24170..24420,24635..25001)  
/gene="MAL3P4.21"  
/note="predicted using hexExon; MAL3P4.21 (PFC0460w),  
Hypothetical protein, len: 206 aa"  
/codon\_start=1  
/protein\_id="CAA15611.1"  
/db\_xref="PID:g1332562"  
/db\_xref="PID:g13764018"  
/db\_xref="GI:3764018"  
/db\_xref="SPTREMBL:O77356"

gene  
CDS

complement(26011..28044)  
/gene="MAL3P4.20"  
complement(Join(26011..27209,27372..27599,27795..28044))  
/gene="MAL3P4.20"  
/note="predicted using hexExon; MAL3P4.20 (PFC0465c),  
Hypothetical protein, len: 559 aa"  
/codon\_start=1  
/protein\_id="CAA15610.1"  
/db\_xref="PID:g1332561"

Query Match 1.2% Score 22; DB 36; Length 113899;  
Best Local Similarity 100.0%; Pred. No. 0.4;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1232 aacacagagagagagcattca 1253  
Db 90303 AACACAGAGAGGACATCTTCA 90282

Search completed: November 20, 1999, 15:27:56  
Job time: 669 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: November 20, 1999, 15:06:39 ; Search time 505.02 seconds  
(without alignments)

Title: US-09-126-945-1  
Perfect score: 1894  
Sequence: 1 gctcgactctccagcac.....ataagatactagaactg 1894  
Scoring table: IDENTITY\_NUCDX

Searched: 679419 seqs, 1590154680 residues  
Database :

GenEmbl:\*

1:	gb_da1:*
2:	gb_da2:*
3:	gb_om:*
4:	gb_ov:*
5:	gb_pat:*
6:	gb_ph:*
7:	gb_p11:*
8:	gb_p12:*
9:	gb_p13:*
10:	gb_p14:*
11:	gb_p15:*
12:	gb_p16:*
13:	gb_p17:*
14:	gb_p18:*
15:	gb_p19:*
16:	gb_p20:*
17:	gb_p21:*
18:	gb_p22:*
19:	gb_p23:*
20:	gb_p24:*
21:	gb_p25:*
22:	gb_p26:*
23:	gb_p27:*
24:	gb_p28:*
25:	gb_p29:*
26:	gb_p30:*
27:	gb_p31:*
28:	gb_p32:*
29:	gb_p33:*
30:	gb_p34:*
31:	gb_p35:*
32:	gb_p36:*
33:	gb_p37:*
34:	gb_p38:*
35:	gb_p39:*
36:	gb_p40:*
37:	gb_p41:*
38:	gb_p42:*
39:	gb_p43:*
40:	gb_p44:*
41:	gb_p45:*
42:	gb_p46:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	
1	1894	100.0	1894	11	AF071538	AF071538 Homo sapi	
2	751.2	39.7	1704	12	AB019436	AB019436 Mus musc	
3	150.8	8.0	362	36	DRODERS4A	M88474 Drosophila	
c	4	147.6	7.8	60056	35	AC007818	AC007818 Drosophila
5	147.6	7.8	1832	37	AF106964	AF106964 Strongyl	
6	99.2	5.2	6479	36	DROE74A	M37082 Drosophila	
7	99.2	5.2	5315	36	DROE74B	M37083 Drosophila	
8	97.6	5.2	4800	36	DM74E	X15087 Drosophila	
9	86.4	4.6	2695	9	AB016193	AB016193 Homo sapi	
10	86.4	4.6	2266	9	HUMELK1A	M25269 Homo sapien	

Query Match	100.0%; Score 1894; DB 11; Length 1894;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1894; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
0Y 1 gtcgacttcctccacacacattcctcgtcacttcgacgtgtccaaacatgccccacagacc 60	416..1423
	/gene="PDEF"
	/note="prostate epithelium-specific"
	/codon_start=-1
	/product="Ets transcription factor PDEF"
	/protein_id="AAC95296.1"
	/db_xref="PID:g4007418"
	/db_xref="GI:4007418"
	/translation="MCSASGLSSVSPPSHLLPPDVTSTRTGLKAAAGVGLRRDS
	PSPAPPEGLSAFVLYSTYPMKLPEDSSNAAPKAGSSREPEPEEQCVIUSQAPA
	GSIDLVPGLITTEHSLSDVOYSMVGELVDITACKLKLWTADPDMSNVQMLP
	WTEQRIHLPKMGAKFELAEKLCAMSEEDFRRSPGGVLAHLADIKMSAAMKER
	TSQALHYCASTSEESSTGDEVDSSCSGDFILFMQFLKELLPLHSGYGRIRVLNTER
	GIRKIEDSAQVARIAMGIRKNRPMANYDKLSRSIROXYKKIIRKPIQSLVYQVFHP
Db 1 gtcgacttcctccacacacattcctcgtcacttcgacgtgtccaaacatgccccacagacc 60	
0Y 61 agtcctccaaagcctctgtcccaagctccctcgaagccctcagtggtgaccttgcaagtg 120	
Db 61 AGTCCTCCAAGCCGCTGCGCAGCTCCCTCGAAGCCCTTCAGGTGTGGCCCTTCGACAGCTG 120	
0Y 121 ccagcagcagcagccttggtgctgctgggtggtgagatcccttaagcagcagcagccttgagac 180	
Db 121 CCAAGCAGCAGCCCTTGGCTGGGGGCTATGGGGACTCCCTACAGGACAGCAGCCCTTGAGACC 180	

78  
79

QY	241	gcaacacagccctctctggccctctgttctccgttgcccccaagatccttgcttgaagaaacgc	300
Dp	241	gccacacagccctctctggccctctgttctccgttgcccccaagatccttgcttgaagaaacgc	300
QY	301	caatggcccttaagtctgccacaacctctcccgccctgaagtctgacatgcagacaag	360
Dp	301	caatggcccttaagtctgccacaacctctcccgccctgaagtctgacatgcagacaag	360
QY	361	ctccctctgggacacaggacagctaaacagacacagccgcacagcccaaacagacagcggaatgg	420
Dp	361	ctccctctgggacacaggacagctaaacagacacagccgcacagcccaaacagacagcggaatgg	420
QY	421	cagcgccagcccggtctctgaagcagctatccccagacacctctctgtgcccccgacac	480
Dp	421	cagcgccagcccggtctctgaagcagctatccccagacacctctctgtgcccccgacac	480
QY	481	ggtctctgcgagacagcgtcttgagaagcgccgacagcgcgggagcgtggtctgtgaagacagga	540
Dp	481	ggtctctgcgagacagcgtcttgagaagcgccgacagcgcgggagcgtggtctgtgaagacagga	540
QY	541	ctggagctccagctccaccccgccacgcgcgcgacagcgacgtctgtctgcctcttacctctcta	600
Dp	541	ctggagctccagctccaccccgccacgcgcgcgacagcgacgtctgtctgcctcttacctctcta	600
QY	601	ctttgacacagctgctaccccttgaaagacagcagctgggacgacaaagcccttgggccagacag	660
Dp	601	ctttgacacagctgctaccccttgaaagacagcagctgggacgacaaagcccttgggccagacag	660
QY	661	tcggggaaggacacacttgaagagaccttgagcagctggccggtcatcttgacagccaaagccccagc	720
Dp	661	tcggggaaggacacacttgaagagaccttgagcagctggccggtcatcttgacagccaaagccccagc	720
QY	721	ggcagagccctggacattggtgtcccgccggcgtgaaccttggaaagagcaactcgtctgaagacagt	780
Dp	721	ggcagagccctggacattggtgtcccgccggcgtgaaccttggaaagagcaactcgtctgaagacagt	780



[illegible]

OY	1238	gagaaggcattcttcaaatgtgagcctcagccaggctggccggctgtgggcatccgc	1397
Db	1154	GAGAAAGCATTCTTCAAAATTTGAGGACTAGCAGAGGTGGCCCGCTGTGGGGTGTGC	1213
OY	1298	aagaaccgtcccgccatcgtgaactacagacagaagcgcgcgtccatccgcgcagatataaag	1357
Db	1214	AAGAAACGGGCGAGCGATGAGACTTGTGATTAACCTTAAGCCGCTCATCCGCGCAATATTACAAG	1273
OY	1358	aagggcattcatccgggaagccagacatctccagcgccctgctataccagttcgtgacccc	1417
Db	1274	AAGGGCATCATGTGTAAACCCGCATCTCTCAGCCCTTGTCTCAATATTGTGCATCA	1333
OY	1418	atccgagtgac-----tgagccagagcgctgaacccgcgcctcaaggagcctctctctgcct	1473
Db	1334	GTCGTAGAGGCCACAGAGACCCAGAGGCGCTTCAACACTGCCCCAGGACCCACTCTCTGTTG	1393
OY	1474	gcacctgcc	1481
Db	1394	GCCCTGGTC	1401

RESULT	3
LOCUS	DRODETSA4
DEFINITION	362 bp DNA INV 05-AUG-1992
ACCESSION	DRODETSA4
NID	Drosophila melanogaster D-ets-4 DNA binding domain protein gene, 3 end.
VERSION	M88474
KEYWORDS	9157195
SOURCE	M88474.1 GI:157195
ORGANISM	DNA-binding domain protein, Drosophila melanogaster DNA. Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 362)
REFERENCE	Chen,T., Buntling,M., Karin,F.D. and Thummel,C.S.
AUTHORS	Isolation and characterization of five Drosophila genes that encode an ets-related DNA binding domain
TITLE	Dev. Biol. 151, 176-191 (1992)
JOURNAL	92249640
MEDLINE	

```

/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/map="Chromosome 3R 98B"
1. .345
gene

```

```

/db_xref="FlyBase:FBgn0005659"
1. .345
CDS

```

```

/gene="Fts98B"
/codon_start=1
/product="D-ets-4 DNA binding domain protein"
/protein_id="AA28451.1"
/db_xref="FID:9157196"
/db_xref="GI:157196"
/translation="TMSNGSTATVKNRPNNGSRGGSGLHMOFLKELLASPOVNGT
AIRNIDSGKGFLEDSVRVAKLGRKRNPAMNYDLSRSIROYKKGIINKKTERSQ
RLVYQFHPYSO"
BASE COUNT      90 a      99 c      113 g      60 t
ORIGIN
Chromosome 3R, 98B.

```

	Query Match	8.0%	Score 150.8;	DB 36;	Length 362;
	Best Local Similarity	72.9%;	Pred. No. 1,7e-17;		
	Matches 194; Conservative	0;	Mismatches 72;	Indels 0;	Gaps 0;
OY	1157 cccatcaccctgtggcgagtttcctccaagaagtgtgctaactcaaggccccacagcatatgcgcgc				1216
I					

```

Db 73 CACATCCACCTGTGCGAGTTCTCTCAAGAGCTGCTGCTTCCGCCAGGTGACGGCACA 132
Oy 1217 tccatgaagtgctcacaagaagagagcattctcaaaattcagaccagccagctg 1276
    ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 133 GCCATCCGGTGTGATCCAGCAAGCAAGGCGATCTTCAAGATCGAGGATCTGGTGAAGG 192
Oy 1277 gccgcgctgtggygcattccgcagaacccgtccgcgcattgaactacagacagctgagccgc 1336
    ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||||
Db 193 GCCAAGCTGTGGGGTGGCGGCAAGAACCGACGGCGATGACTATGATATGATTGTCCCGC 252
Oy 1337 tccatccgcagatattacagaagaagcattccgcgaagccagacacattccagcgcc 1396
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 253 TCCATCAGGACGACTACTACAGAAAGGAAATCATGAAAGACGAGAGGTACACAGCGGCTG 312
Oy 1397 gtctacagatctgtgcagccacatg 1422
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 313 GTCTACGAGTGTCTCCATCCCTACAG 338

RESULT 4
AC007818/c
LOCUS DEFINITION AC007818 60056 bp DNA HTG 14-JUN-1999
Drosophila melanogaster chromosome 3 clone BACR02M05 (D785) RPCI-98
02.M.5 map 98B-98B strain Y: cn bw sp, WORKING DRAFT SEQUENCE, 20
unordered pieces.
ACCESSION AC007818 95053146
VERSION AC007818.1 GI:5053146
KEYWORDS HTG; HTGS_PHASE1.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 60056)
AUTHORS Celisner,S.E., Abdayani,A., Arcana,T.T., Baxter,E., Blazey,R.G.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomutan,M.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacled,J.M., Park,S.,
Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Solt,E.,
Svitskys,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
Rubin,G.M.
Sequencing of Drosophila melanogaster
Unpublished
2 (bases 1 to 60056)
AUTHORS Celisner,S.E., Abdayani,A., Arcana,T.T., Baxter,E., Blazey,R.G.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
Kim,E., Lee,B., Lewis,S., Li,P., Lomutan,M.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacled,J.M., Park,S.,
Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Solt,E.,
Svitskys,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
Rubin,G.M.
Direct Submission
Submitted (14-JUN-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive web site (http://www.fruitfly.org/sequence/) or send email
to bdg@fruitfly.berkeley.edu. All contigs in this submission meet
the following cutoffs: length >= 400 bases, phrap computed error
rate <= 1/10.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

* 1 893: contig of 893 bp in length
* 894 973: gap of unknown length
* 974 1964: contig of 991 bp in length
* 1965 2044: gap of unknown length
* 2045 3652: contig of 1608 bp in length
* 3653 3732: gap of unknown length
* 3733 5272: contig of 1540 bp in length
* 5273 5352: gap of unknown length
* 5353 6476: contig of 1124 bp in length
* 6477 6557: gap of unknown length
* 6557 8670: contig of 2113 bp in length
* 8670 8750: gap of unknown length
* 8750 11281: contig of 2532 bp in length
* 11282 11361: gap of unknown length
* 11362 13275: contig of 1914 bp in length
* 13276 13355: gap of unknown length
* 13356 16154: contig of 2799 bp in length
* 16155 16234: gap of unknown length
* 16235 18409: contig of 2175 bp in length
* 18410 18489: gap of unknown length
* 18490 20047: contig of 1558 bp in length
* 20048 20127: gap of unknown length
* 20128 21987: contig of 1860 bp in length
* 21988 22067: gap of unknown length
* 22068 24914: contig of 2847 bp in length
* 24915 24994: gap of unknown length
* 24995 28786: contig of 3792 bp in length
* 28787 28866: gap of unknown length
* 28867 32920: contig of 4054 bp in length
* 32921 33000: gap of unknown length
* 33001 36522: contig of 3522 bp in length
* 36523 36602: gap of unknown length
* 36603 41890: contig of 5287 bp in length
* 41890 41969: gap of unknown length
* 41970 45112: contig of 4542 bp in length
* 45112 46512: gap of unknown length
* 46512 51474: contig of 4883 bp in length
* 51475 51554: gap of unknown length
* 51555 60056: contig of 8502 bp in length.
* Location/Qualifiers
* 1..60056
  /organism="Drosophila melanogaster"
  /strain="Y: cn bw sp"
  /db_xref="taxon:7227"
  /clone="BACR02M05 (D785) RPCI-98 02.M.5"
  /chromosome="3"
  /clone_11b="RPCI-98 (Roswell Park Cancer Institute
  /Drosophila melanogaster BAC library, partial EcoRI in
  pBACE3.6"
  /map="98B-98B"
BASE COUNT 15556 a 13397 c 13635 g 15945 t 1523 others
ORIGIN
Query Match 7.8% Score 147.6; DB 35; Length 60056;
Best Local Similarity 72.2%; Pred. No. 2,16-17;
Matches 192; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
Oy 1157 cccatccacccgtgtgcattctctcaagaagttgtcattcccaagccccaagctatgcccgc 1216
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 54077 CACATCCACCTGTGCGAGTTCTCTCAAGAGCTGCTGCTTCCGCCAGGTGACGGCACA 54018
Oy 1217 tccatgaagtgctcacaagaagagagcattctcaaaattcagaccagccagctg 1276
    ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||||
Db 54017 GCCATCCGGTGTGATCCAGCAAGCAAGGCGATCTTCAAGATCGAGGATCTGGTGAAGG 53958
Oy 1277 gccgcgctgtggygcattccgcagaacccgtccgcgcattgaactacagacagctgagccgc 1336
    ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||||
Db 53957 GCCAAGCTGTGGGGTGGCGGCAAGAACCGACGGCGATGACTATGATATGATTGTCCCGC 53898
Oy 1337 tccatccgcagatattacagaagaagcattccgcgaagccagacacattccagcgcc 1396
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 53897 TCCATCAGGACGACTACTACAGAAAGGAAATCATGAAAGACGAGAGGTACACAGCGGCTG 53858

```





```

polyA_signal
/gene="Elip74EF"
/db_xref="FlyBase:FBN0000567"
BASE COUNT      1465 a      1423 c      1304 g      1123 t
ORIGIN
Query Match          5.2% Score 99.2; DB 36; Length 5315;
Best Local Similarity 64.5%; Pred. No. 9, 2e-09;
Matches 165; Conservative 0; Mismatches 88; Indels 3; Gaps 1;

QY 1158 ccattcaccttggcagttccctcaaggatgtgctaactaaagcccacagctatggcgct 1217
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3408 CCACCTATCTGTGGAGATTCTCCTCAACAGTCCTCCAGATCGGAAATACGTGCCTT 3467

OY 1218 tcatttagtgctcaacaagaagaaggagcgcctctcaaatgtaggaactcagccagctgg 1277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3468 TCATCATMACTGACGACGACCAGGAAGAGGCCGTCTTCMACTGTGTGACTCGAAGCTGTGT 3527

OY 1278 ccgcgctgtgggagcatccgcgaagaaccgctccgcgcatagaactagcagtcgacct 1337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3528 CCCGCTGTGGGATGATGACAAGAACAAGCCGAGATGACTACGATGAGATGGGACAGG 3587

OY 1338 ccatacgcagtatatacagaagaaggatcctaccggaaagccagacatctcccaagccccg 1397
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3588 CGCTGAGGTACTACTACCAAGCGCGCATCTAGCCCAAGGTGGA---TGCCACAGCGCTGG 3644

QY 1398 tctacagttcgtcca 1413
Db 3645 TCTACACGTTCTGTGGA 3660

RESULT 8
DW74E DW74E 4800 bp DNA INV 12-SEP-1993
LOCUS Drosophila melanogaster 74E gene.
DEFINITION X15087
ACCESSION 67513
NID X15087.1 GI:7513
VERSION 74E gene.
KEYWORDS fruit fly.
SOURCE Drosophila melanogaster
ORGANISM Eukaryote; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
Drosophilidae; Drosophila.
1 (bases 1 to 4800)
REFERENCE Pongs,O.
AUTHORS Direct Submission
TITLE Submitted (26-APR-1989) Pongs O., Lehrstuhl fuer Biochemie,
JOURNAL Abteilung Chemie, Ruhr Universitaet Bochum, Universitaestr 150,
D-4630 Bochum, F R G
2 (bases 1 to 4800)
AUTHORS Junknecht,R., Taube,W., Ludecke,H.J. and Pongs,O.
TITLE Characterization of a putative transcription factor gene expressed
REFERENCE In the 20-OH-ecdysone inducible puff 74EF in Drosophila
melanogaster
JOURNAL Nucleic Acids Res. 17 (12), 4455-4464 (1989)
MEDLINE 89315191
COMMENT Data kindly reviewed (10-Jul-1989) by Pongs O.
FEATURES
source Location/Qualifiers
1. .4800
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="JOT4, JOT5, 74E1, 74E2"
/map="chromosome 3"
1. .700
/gene="Elip74EF"
misc_feature

```

[illegible]

polyA_site	4800	/db_xref=Flybase:FBgn00000567
/gene="E1p74EE"		
/note="polyA_site"		
/db_xref=Flybase:FBgn00000567		
BASE COUNT	1381 a 1276 c 1142 g 1001 t	
ORIGIN		

Query Match	5.2%	Score 97.6	DB 36	Length 4800
Best Local Similarity	64.1%	Pred. No. 1.8e-08		
Matches 164	Conservative	0	Mismatches 89	Indels 3
			Gaps	1
QY 1158	ccatccaccctctgacagcttcctcaagaagtgtactatcaagcccccacagctatgcygcct	1217		
Db 2880	CCACCTATCTGTGGAGAGTTCCTCTCAACAGCTCTCCAGATCCGGAAATACGTCCGGCTT	2949		
QY 1218	tcaatgaatgctcacaacaagaagaagggcatcttcacaatatgagagctcagccagcttg	1277		
Db 2950	TCATCAATGTGACGACACCGGGGAGAGGGCGTTCCTAACACTGTGTGACTCGAAGCTGTGT	3009		
QY 1278	ccggagctctggtgagatccgcgaagaacgcgtccgcacatgaactacgacaagctgagcct	1337		
Db 3010	CCCCGCTCTGTGGGCGCTGACACAGAACGACCGGACATGCAATACGAGACGATGGCGCAGG	3069		
QY 1338	gcattccgcagatathacaagaagaagggcatcatccggaagccagacatctcccaagcgcttg	1397		
Db 3070	CGCTGAGGTACTACTACACGACGCGCGCATACGTGGCCAGAGTGA---TGGCCAGGCGCTGG	3126		
QY 1398	tctaacagctgtgtgca 1413			
Db 3127	TCTACCATTTCTGTGA 3142			

RESULT	9				
AB016193					
LOCUS	AB016193	2695	bp	mRNA	PR1
					07-APR-1999

DEFINITION	Homo sapiens Elkl mRNA, complete cds.
ACCESSION	AB016193
NID	94126581
VERSION	AB016193.1 GI:4126581
KEYWORDS	Elkl.
SOURCE	Homo sapiens adult hippocampus cDNA to mRNA, clone:pekl1 4-1.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (sites)
TITLE	Yamauchi,T., Toko,M., Suga,M., Hatakeyama,T. and Isobe,M. Structural organization of the human Elkl gene and its processed pseudogene Elk2
JOURNAL	DNA Res. 6 (1), 21-27 (1995)
MEDLINE	99245057
REFERENCE	2 (bases 1 to 2695)
AUTHORS	Isobe,M., Yamauchi,T. and Toko,M.
TITLE	Direct Submission
JOURNAL	Submitted (10-Jul-1998) to the DDBJ/EMBL/Genbank databases. Masaaki Isobe, Toyama University, Faculty of Engineering, Department of Materials and Biosystem Engineering, 3190 Gofuku, Toyama, Toyama 930-8555, Japan (E-mail:isobe@eng.toyama-u.ac.jp, Tel:+81-764-45-6872, Fax:+81-764-45-6874)
COMMENT	Sequence updated (21-Jul-1998).
FEATURES	location/Qualifiers
SOURCE	1..2695

BASE	COORDNT	541 a	811 c	721 g	622 +
exon	1..120	1..120			
exon	/number=1	67..310			
gene	/number=3	101..1387			
CDS	/gene="Elk1"	101..1387			
	/gene="Elk1"				
	/note="transcription factor"				
	/codon_start=1				
	/protein_id="BAA36616.1"				
	/db_xref="PID:d1037603"				
	/db_xref="PID:q4126582"				
	/db_xref="GI:4126582"				
	/translation="MDPSVTLMQFLIQLLRDEONGHIIISWTSRDGGEFLVDAEEVAR LNGLRNRKTRNNMNDKLSRALRTYDKNIIIRKVSQGVFVYKFSVSPVACSTEDCPQ PEVSYSTNPENVAPPAIHAAPGDVTYSGKPGTGRAGAMAPGGIARSSRNREYAGLPS VTITISLOPPPPHPPAVVLPNNAAPAGAAPSGSRTSPSLAECLAEAEAGLPIQ VLTTPSEAPNLKSEELNVEPGLRALPPEVKVEGPKKEELVAGEIRGFVETTKAEPEV PQEGGVPARLPAYVMDTACGAGHAASPEISQPOKGRKRDLELPLSPISLIGCGPE RTPGSGSGVLAQVAPGALPRLPLPTHTLTPLVLTLSLSPISIHFWNTSLP IAPRSFAK LSFGPPSSQAQVHPIPSISYVDGLSTPVIISPGQPK"				
exon	311..754				
	/gene="Elk1"				
	/number=4				
exon	755..1186				
	/gene="Elk1"				
	/number=5				
exon	1187..1288				
	/gene="Elk1"				
	/number=6				
exon	1289..2695				
	/number=7				
	1385..2695				
3'UTR	541 a	811 c	721 g	622 +	

ORIGIN	Query Match
	4.68; Score 86.4; DB 9; Length 2695;

Best Local Similarity 60.5%; Pred. No. 1.8e-06;  
Matches 179; Conservative 0; Mismatches 111; Indels 6; Gaps 2;

QY 1166 ctgttgacagttcctcaaggagttgtactcaagcccaagcattgacgcttattag 1225  
 Db 119 CTGTGGGACGTTTCTGCTGCAGCTCTGAGAGAG---CAAGGCATGGCCACATCATCTCC 175  
 QY 1226 tggctcaacaaggaaaggacatcttcaaaattgaggactcaagcccgctgagccgctg 1285  
 Db 176 TGGACTTCACGGAGTGGTGTGAATTCAGCTGTGATGACAGAGAGAGTGGCCGCTG 235  
 QY 1286 tgggacatccgcaagaccgctcccgcatgaactacagacagctgagccgctccatccgc 1345  
 Db 236 TGGGACTACGCAAGAACAGACACATGATTAACAGCAAGCTCAGCCGGCCTTGCGG 295  
 QY 1346 cagtatatacaagaaggacatcatcgcgaagcagacatctccagcgctcgtctaccag 1405  
 Db 296 TACTACTATGACAGAACATCATCTCCGCAAG---GTGAGCGGCCAGAAAGTTCTGTACAG 352  
 QY 1406 ttctgtgaccccatctgagtgctgctgcccagggcctgaaacccgctccaggggccc 1461  
 Db 353 TTGTGTCTTACCTCGTAGGTGCGAGGGTCTCTCCACTGAGACTCCCGCCCAAGCC 408

## RESULT 10

HUMELK1A 2266 bp mRNA PRI 08-NOV-1994  
 LOCUS Homo sapiens tyrosine kinase (ELK1) oncogene mRNA, complete cds.  
 DEFINITION M25269  
 ACCESSION 9538208  
 NID M25269.1 GI:538208  
 VERSION ETS1 gene; oncogene: tyrosine kinase.  
 KEYWORDS Homo sapiens cDNA to mRNA.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2266)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 AUTHORS Rao,Y.N., Huebner,K., Isobe,M., ar-Ruehnl,A., Croce,C.M. and Reddy,E.S.

TITLE elk, tissue-specific ets-related genes on chromosomes X and 14 near translocation breakpoints  
 JOURNAL Science 244 (4900), 66-70 (1989)  
 MEDLINE 89203250

COMMENT On Sep 15, 1994 this sequence version replaced gi:341319.

## FEATURES

source  
 1..2266  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /cell\_line="COLO 320"  
 /clone="lambda-11"  
 /map="Xp22.1-p11"  
 316..1602  
 /gene="ELK1"  
 /gene="ELK1"  
 316..1602  
 /codon\_start=1  
 /db\_xref="GDB:600-119-867"  
 /product="tyrosine kinase"  
 /protein\_id="AA52384.1"  
 /db\_xref="PID:G538209"  
 /db\_xref="GI:538209"  
 /translation="MDPSVTLMOFLQLREOGNGHIIHSITSDGGEFLKYDAEVAR  
 LMGIRKNTMNYDKLSBALRYVDNRIIRKYSQGFYKFGVSPVACSTEDCDPO  
 PEVSTSTNPNVPAIHAHAGTIVSGKPTPGAGAGGGLARSNRNRYMAGLS  
 TFIQLQPPPHPRPRAVLPNAPAGAAAPSGSRSTSPSLEALEEAGLPLQ  
 VILTPPEAPNLKSEELNVEPGLRALPEYKVGPREELVAGERGFVEPTAAEYV  
 PDEGVPAFLPVMVDTAGACGHAASPEVISOPOKGRKRDLELPSLLGSGPE  
 RTSGSGSLGAPGALPRLPLTPTHTLPTPLTSPSLPSIHFMSTLPSIADRSAPK  
 LSPFPSSGSAOVHIPSISYDGLSTPVLTPSPGRK"

## BASE COUNT

457 a 731 c 605 g 473 t

Query Match 4.6%; Score 86.4; DB 9; Length 2266;  
 Best Local Similarity 60.5%; Pred. No. 1.9e-06;  
 Matches 179; Conservative 0; Mismatches 111; Indels 6; Gaps 2;

QY 1166 ctgttgacagttcctcaaggagttgtactcaagcccaagcattgacgcttattag 1225  
 Db 334 CTGTGGGACGTTTCTGCTGCAGCTCTGAGAGAG---CAAGGCATGGCCACATCATCTCC 390  
 QY 1226 tggctcaacaaggaaaggacatcttcaaaattgaggactcaagcccgctgagccgctg 1285  
 Db 391 TGGACTTCACGGAGTGGTGTGAATTCAGCTGTGATGACAGAGAGAGTGGCCGCTG 450  
 QY 1286 tgggacatccgcaagaccgctcccgcatgaactacagacagctgagccgctccatccgc 1345  
 Db 451 TGGGACTACGCAAGAACAGACACATGATTAACAGCAAGCTCAGCCGGCCTTGCGG 510  
 QY 1346 cagtatatacaagaaggacatcatcgcgaagcagacatctccagcgctcgtctaccag 1405  
 Db 511 TACTACTATGACAGAACATCATCTCCGCAAG---GTGAGCGGCCAGAAAGTTCTGTACAG 567  
 QY 1406 ttctgtgaccccatctgagtgctgctgcccagggcctgaaacccgctccaggggccc 1461  
 Db 568 TTGTGTCTTACCTCGTAGGTGCGAGGGTCTCTCCACTGAGACTCCCGCCCAAGCC 623

## RESULT 11

AF075706 848 bp mRNA VRT 19-DEC-1998  
 LOCUS Gallus gallus ets domain protein (ERM) mRNA, partial cds.  
 DEFINITION AF075706  
 ACCESSION 93869357  
 NID AF075706.1 GI:3869357  
 VERSION AF075706.1 GI:3869357  
 KEYWORDS chicken.  
 SOURCE Gallus gallus  
 ORGANISM Gallus gallus

REFERENCE 1 (bases 1 to 848)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria;  
 Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.  
 AUTHORS Lin,J.H., Salto,T., Anderson,D.J., Lance-Jones,C., Jessell,T.M. and Arber,S.

TITLE functionally related motor neuron pool and muscle sensory afferent subtypes defined by coordinate ETS gene expression  
 JOURNAL Cell 95 (3), 393-407 (1998)  
 MEDLINE 99029846

REFERENCE 2 (bases 1 to 848)  
 Lin,J.H., Salto,T., Anderson,D.J., Lance-Jones,C., Jessell,T.M. and Arber,S.

Direct Submission  
 Submitted (01-JUL-1998) Center for Neurobiology & Behavior,  
 Columbia University, 701 W. 168th St. HHSC 1013, New York, NY  
 10032, USA

## FEATURES

source  
 1..848  
 /organism="Gallus gallus"  
 /db\_xref="taxon:9031"  
 <1..848  
 /gene="ERM"  
 <1..712  
 /gene="ERM"  
 /codon\_start=2  
 /product="ets domain protein"  
 /protein\_id="AAC97201.1"  
 /db\_xref="PID:G3869358"  
 /db\_xref="GI:3869358"  
 /translation="RPQPARPRPHGOSMGKIOEPRDYCIDSEVPNCGSSVYRGCLF  
 PSHDSFSEKRTLFDTQVAPERLEKVAOEPLLYEGEPYORLQWLQWFLVT  
 LDDPANNAFIWMTGKMEFKLIEPERLRNGIONRRAMNVDKLSRLRYEKGKI  
 MOKVAGERVYKFCVDPDALFSMAYPDNRPLAKEPDCHVSEEDTLPLTHFEDSPAY  
 LLEVERGSLPFAEGPAY"

## BASE COUNT

192 a 265 c 245 g 146 t





REFERENCE	3	(bases 1 to 1403)
AUTHORS	Dojima et al.	0

DERACCE, U.  
ATTI D

Direct submission

JOURNAL  
Submitted (13-FEB-1997) O. Delattre, Institut Curie, Inserm U434

26 Rue Dulm, Paris 75231 Cedex 05, Paris 75231 Cedex 05, FRANCE

COMMENT On Feb 15, 1997 this sequence version replaced nt.1834A85

### FEATURES

Location/Qualifiers

**Source**

Организация "Национал-демократический фронт"

Organisms - Homo sapiens

uu\_xref=

/germline

incron 1. .240

/number=2

/evidence=

gene 241 1403

000000="FEV"

79-1403  
JUNE- FEB  
241 1403  
6000

EXON 241. 1403

```
/gene="FEV"
```

Query Match 4.48; Score 83.4; DB 10; Length 1403;

Best Local Similarity 54.7%; Pred. No. 6.9e-06;

Matches	214;	Conservative	0;	Mismatches	166;	Indels	11;	Gaps	2;
---------	------	--------------	----	------------	------	--------	-----	------	----

1160 atccacctgtgcagttcctcaaggagattgactcaagggcgcacagtatggcgccttc

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

[illegible]

STANDARD 311

0v 1220 at+acatccccccccc-----

1220 acacacagggcattccaaatctgaggaccaggtggcc

2

312 GCGTGGGAGGCC-----GTCACGGCGAGTTCAGCTCACGGACCCGGACGAGTGGCG 365

QY 1280 cggctgtggcatccgcaaacctccgccatgaaactaacgaaactgaacacac 1330

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100

[illegible][illegible]

Search completed: November 20, 1999, 15:16:44  
Job time: 605 sec